

# Raw Sequence Listing Error Summary

<u>ERROR DETECTED</u>	<u>SUGGESTED</u>	<u>CORRECTION</u>	<u>SERIAL NUMBER:</u> <u>09/816, 391</u>
<b>ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE</b>			
1 <input checked="" type="checkbox"/> Wrapped Nucleic		The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3, as this will prevent "wrapping".	
2 <input type="checkbox"/> Wrapped Aminos		The amino acid number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3, as this will prevent "wrapping".	
3 <input type="checkbox"/> Incorrect Line Length		The rules require that a line not exceed 72 characters in length. This includes spaces.	
4 <input type="checkbox"/> Misaligned Amino Acid Numbering		The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs between the numbering. It is recommended to delete any tabs and use spacing between the numbers.	
5 <input type="checkbox"/> Non-ASCII		This file was not saved in ASCII (DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text so that it can be processed.	
6 <input type="checkbox"/> Variable Length		Sequence(s) ____ contain n's or Xaa's which represented more than one residue. As per the rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the (ix) feature section that some may be missing.	
7 <input type="checkbox"/> PatentIn ver. 2.0 "bug"		A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequence(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. <b>This applies primarily to the mandatory &lt;220&gt;-&lt;223&gt; sections for Artificial or Unknown sequences.</b>	
8 <input type="checkbox"/> Skipped Sequences (OLD RULES)		Sequence(s) ____ missing. If intentional, please use the following format for each skipped sequence: <b>(2) INFORMATION FOR SEQ ID NO:X:</b> <b>(i) SEQUENCE CHARACTERISTICS:</b> (Do not insert any headings under "SEQUENCE CHARACTERISTICS") <b>(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:</b> This sequence is intentionally skipped  Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).	
9 <input type="checkbox"/> Skipped Sequences (NEW RULES)		Sequence(s) ____ missing. If intentional, please use the following format for each skipped sequence. <210> sequence id number <400> sequence id number 000	
10 <input type="checkbox"/> Use of n's or Xaa's (NEW RULES)		Use of n's and/or Xaa's have been detected in the Sequence Listing. Use of <220> to <223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.	
11 <input type="checkbox"/> Use of "Artificial" (NEW RULES)		Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.	
12 <input type="checkbox"/> Use of <220>Feature (NEW RULES)		Sequence(s) ____ are missing the <220>Feature and associated headings. Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial Sequence" or "Unknown" Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of new Rules)	
13 <input type="checkbox"/> PatentIn ver. 2.0 "bug"		Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other means to copy file to floppy disk.	

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/816,391

DATE: 04/09/2001  
TIME: 09:43:22

Input Set : A:\sequence list.txt  
Output Set: N:\CRF3\04092001\I816391.raw

3 <110> APPLICANT: KYOWA HAKKO KOGYO CO., LTD.  
W--> 4 <120> TITLE OF INVENTION: Anaerobic bacterium as a drug for cancer gene therapy  
W--> 5 <130> FILE REFERENCE: DK04F318(US)  
W--> 6 <140> CURRENT APPLICATION NUMBER:  
C--> 7 <141> CURRENT FILING DATE: 2001-03-26  
8 <150> PRIOR APPLICATION NUMBER: JP 00/287688  
W--> 9 <151> PRIOR FILING DATE: 2000-9-21  
W--> 10 <160> NUMBER OF SEQ ID: 3

Does Not Comply  
Corrected Diskette Needed

pp. 1, 2

#### ERRORED SEQUENCES

W--> 11 <210> SEQ ID NO: 1  
12 <211> LENGTH: 600  
13 <212> TYPE: DNA  
14 <213> ORGANISM: Bifidobacterium longum  
W--> 15 <220> FEATURE:  
16 <221> NAME/KEY: CDS  
17 <222> LOCATION: (193)..(471)  
W--> 18 <400> SEQUENCE: 1  
E--> 19 gctgggcgcg gcgccatga agtggcttga caagcataat cttgtctgat tcgtcttattt  
20 (60) \_\_\_\_\_  
E--> 21 tcaataacctt cggggaaaata gatgtgaaaa cccttataaaa acgcgggttt tcgcagaaaac  
22 (120) \_\_\_\_\_  
E--> 23 atgcgcctagt atcattgtat acaacatgga ctaagcaaaa gtgcttgcct cctgacccaa  
24 (180) \_\_\_\_\_  
E--> 25. gaaggatgt tt atg gca tac aac aag tct gac ctc gtt tcg aag atc gcc  
26 (231) \_\_\_\_\_  
27 Met Ala Tyr Asn Lys Ser Asp Leu Val Ser Lys Ile Ala move to here.  
W--> 28 1 5 10  
E--> 29 cag aag tcc aac ctg acc aag gct cag gcc gag gct gct gtt aac gcc  
30 (279) \_\_\_\_\_  
31 Gln Lys Ser Asn Leu Thr Lys Ala Gln Ala Glu Ala Ala Val Asn Ala  
W--> 32 15 20 25  
E--> 33 ttc cag gat gtg ttc gtc gag gct atg aag tcc ggc gaa ggc ctg aag  
34 (327) \_\_\_\_\_  
35 Phe Gln Asp Val Phe Val Glu Ala Met Lys Ser Gly Glu Gly Leu Lys  
W--> 37 30 35 40 45  
E--> 38 ctc acc ggc ctg ttc tcc gct gag cgc gtc aag cgc ccg gct cgc acc  
39 (375) \_\_\_\_\_  
W--> 40 Leu Thr Gly (Lue) Phe Ser Ala Glu Arg Val Lys Arg Pro Ala Arg Thr  
W--> 41 50 55 60  
E--> 42 ggc cgc aac cgc cgc act ggc gag cag att gac att ccg gct tcc tac  
43 (423) \_\_\_\_\_  
44 Gly Arg Asn Pro Arg Thr Gly Glu Gln Ile Asp Ile Pro Ala Ser Tyr  
W--> 46 65 70 75  
E--> 47 ggc gtt cgt atc tcc gct ggc tcc ctg ctg aag aag gcc gtc acc gag  
(471) \_\_\_\_\_

"Wrapped"  
nucleic acid  
numbers. See  
# 1 on Error  
sheet.  
Summary

Invalid amino acid designator

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/816,391

DATE: 04/09/2001  
TIME: 09:43:22

Input Set : A:\sequence list.txt  
Output Set: N:\CRF3\04092001\I816391.raw

48                   471  
49 Gly Val Arg Ile Ser Ala Gly Ser Leu Leu Lys Lys Ala Val Thr Glu  
50                   80                   85                   90  
W--> 51 tgaccttctg ctcgtacgca ttacttcgag cattactgac gacaaagagacc ccgaccgaga  
E--> 52                   531  
E--> 53 tggtcgggggt ctttttgttg tggtgctgtg acgtgttgtc caaccgtatt attccggact  
54                   591  
E--> 55 agttcagcg  
56                   600

"} "wrapped"  
"nucleic acid  
numbers. See  
#1 on the

Error Summary  
Sheet.

VERIFICATION SUMMARY DATE: 04/09/2001  
PATENT APPLICATION: US/09/816,391 TIME: 09:43:23

Input Set : A:\sequence list.txt  
Output Set: N:\CRF3\04092001\I816391.raw

L:4 M:283 W: Missing Blank Line separator, <120> field identifier  
L:5 M:283 W: Missing Blank Line separator, <130> field identifier  
L:6 M:283 W: Missing Blank Line separator, <140> field identifier  
L:7 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:9 M:256 W: Invalid Numeric Header Field, Wrong Prior FILING DATE:YYYY-MM-DD  
L:10 M:283 W: Missing Blank Line separator, <160> field identifier  
L:11 M:283 W: Missing Blank Line separator, <210> field identifier  
L:15 M:283 W: Missing Blank Line separator, <220> field identifier  
L:18 M:283 W: Missing Blank Line separator, <400> field identifier  
L:19 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:60 SEQ:1  
M:254 Repeated in SeqNo=1  
L:28 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:32 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:37 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:40 M:334 W: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:1  
L:41 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:46 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:50 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
L:61 M:283 W: Missing Blank Line separator, <220> field identifier  
L:63 M:283 W: Missing Blank Line separator, <400> field identifier  
L:69 M:283 W: Missing Blank Line separator, <220> field identifier  
L:71 M:283 W: Missing Blank Line separator, <400> field identifier